Title	離散型entropy power inequalityについて (非加法性の数理と情報 : 凸解析との接点)
Author(s)	柳, 研二郎
Citation	数理解析研究所講究録 (2010), 1683: 111-120
Issue Date	2010-04
URL	http://hdl.handle.net/2433/141412
Right	
Type	Departmental Bulletin Paper
Textversion	publisher

数理解析研究所講究録
第 1683 巻 2010 年 111-120

離散型 entropy power inequality について
(On discrete entropy power inequality)

山口大学大学院・理工学研究科 柳 研二郎 (Kenjiro Yanagi)*
(Graduate School of Science and Engineering, Yamaguchi University)

Key Words: information theory, entropy power inequality

MSC(2000): 94A17

1 Entropy Power 不等式

連続型の Entropy Power 不等式

$$e^{2h(X)} + e^{2h(Y)} \leq e^{2h(X+Y)} \tag{1}$$

は Shannon [6] が初めて示したもので, その不等式が成り立つことはよく知られている. 後に, 関連した不等式が示されたが, 離散型の Entropy Power 不等式

$$e^{2H(X_m)} + e^{2H(X_n)} \leq e^{2H(X_m+X_n)} = e^{2H(X_{m+n})} \tag{2}$$

は一般には成り立たないことが知られている. 離散型の Entropy Power 不等式が成り立つ特別な場合の考察を行う.

2 超加法性

Definition 1 関数 $n \curvearrowright Y_n$ が以下をみたすとき, 超加法性をもつという.

$$^{\forall}m, n \quad Y_{m+n} \geq Y_m + Y_n.$$

超加法性の十分条件は以下の結果から与えられる.

Proposition 1 $\frac{Y_n}{n}$ が n 関して増加であるとき Y_n は超加法性をもつ.
また Y_n が超加法性をもつならば $\lim_{n\to\infty} \frac{Y_n}{n} = \sup \frac{Y_n}{n}$ が存在する.

*This research was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (B), 18300003 and (C), 20540175

Proof. 任意の m, n $(m \geq n)$ に対して次が成り立つ.

$$\frac{Y_{m+n}}{m+n} \geq \frac{Y_m}{m}.$$

すなわち

$$Y_{m+n} \geq Y_m + \frac{n}{m} Y_m.$$

$m \geq n$ より

$$\frac{Y_m}{m} \geq \frac{Y_n}{n}.$$

よって

$$Y_{m+n} \geq Y_m + Y_n.$$

□

$$Y_n = e^{2H(X_n)} \tag{3}$$

が超加法性をもつことを示すためには, $n \curvearrowright Y_n$ が n に関して増加であることを示せば十分である.

3 情報理論の不等式

Definition 2 独立な二項分布 $X_n \sim B(n, p)$ とベルヌイ分布 $B \sim Ber(p)$ に対して次の確率変数を定義する.

$$X_{n+1} = X_n + B \tag{4}$$

このとき,

$$X_{n+1} = X_n + B = \begin{cases} X_n & B = 0 \text{ すなわち, 確率 } q \text{ のとき} \\ X_n + 1 & B = 1 \text{ すなわち, 確率 } p \text{ のとき} \end{cases}$$

となる. また, 確率分布は次のような関係がある.

$$P_{X_{n+1}} = P_{X_n} * P_B = qP_{X_n} + pP_{X_n+1} \tag{5}$$

それぞれの確率分布は次のようになる.

$$\begin{aligned} Pr(X_n = k) &= \binom{n}{k} p^k q^{n-k}, \quad k = 0, 1, \cdots, n \qquad (6) \\ Pr(X_n = n+1) &= 0 \end{aligned}$$

また,

$$
\begin{aligned}
Pr(X_n+1=k) &= \binom{n}{k-1} p^{k-1} q^{n-k+1}, \quad k=1,\cdots,n+1 \qquad (7)\\
Pr(X_n+1=0) &= 0
\end{aligned}
$$

さらに, (5),(6),(7) より,

$$Pr(X_{n+1}=k) = \binom{n+1}{k} p^k q^{n+1-k}, \quad k=0,1,\cdots,n+1$$

となる. すなわち

$$X_{n+1} \sim B(n+1,p).$$

Lemma 1

$$
\begin{aligned}
& H(P_{X_{n+1}}) \\
=\ & pH(P_{X_n+1}) + qH(P_{X_n}) \\
& +pD(P_{X_n+1}||P_{X_{n+1}}) + qD(P_{X_n}||P_{X_{n+1}}) \qquad (8)
\end{aligned}
$$

Proof. 以下のエントロピー及び相互情報量を用いて (8) を示す.

$$
\begin{aligned}
H(P_{X_{n+1}}) &= -\sum_{i=0}^{n+1} \binom{n+1}{i} p^i q^{n+1-i} \log \binom{n+1}{i} p^i q^{n+1-i} \\
H(P_{X_n+1}) &= -\sum_{i=1}^{n+1} \binom{n}{i-1} p^{i-1} q^{n-i+1} \log \binom{n}{i-1} p^{i-1} q^{n-i+1} \\
H(P_{X_n}) &= -\sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} \log \binom{n}{i} p^i q^{n-i} \\
D(P_{X_n+1}||P_{X_{n+1}}) &= \sum_{i=1}^{n+1} \binom{n}{i-1} p^{i-1} q^{n-i+1} \log \frac{\binom{n}{i-1} p^{i-1} q^{n-i+1}}{\binom{n+1}{i} p^i q^{n+1-i}} \\
&= \sum_{i=1}^{n+1} \binom{n}{i-1} p^{i-1} q^{n-i+1} \log \frac{i}{p(n+1)} \\
D(P_{X_n}||P_{X_{n+1}}) &= \sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} \log \frac{\binom{n}{i} p^i q^{n-i}}{\binom{n+1}{i} p^i q^{n+1-i}} \\
&= \sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} \log \frac{n+1-i}{q(n+1)}
\end{aligned}
$$

ここで (8) の右辺は,

$$
\begin{aligned}
& pH(P_{X_n+1}) + qH(P_{X_n}) + pD(P_{X_n+1}||P_{X_{n+1}}) + qD(P_{X_n}||P_{X_{n+1}}) \\
= \; & p\left(-\sum_{i=1}^{n+1}\binom{n}{i-1}p^{i-1}q^{n-i+1}\log\binom{n}{i-1}p^{i-1}q^{n-i+1}\right) \\
& +q\left(-\sum_{i=0}^{n}\binom{n}{i}p^{i}q^{n-i}\log\binom{n}{i}p^{i}q^{n-i}\right) \\
& +p\left(\sum_{i=1}^{n+1}\binom{n}{i-1}p^{i-1}q^{n-i+1}\log\frac{i}{p(n+1)}\right) \\
& +q\left(\sum_{i=0}^{n}\binom{n}{i}p^{i}q^{n-i}\log\frac{n+1-i}{q(n+1)}\right) \\
= \; & p\left(\sum_{i=1}^{n+1}\binom{n}{i-1}p^{i-1}q^{n-i+1}\log\frac{i}{p(n+1)\binom{n}{i}p^{i-1}q^{n-i+1}}\right) \\
& +q\left(\sum_{i=0}^{n}\binom{n}{i}p^{i}q^{n-i}\log\frac{n+1-i}{q(n+1)\binom{n}{i}p^{i}q^{n-i}}\right) \\
= \; & p\left(\sum_{i=1}^{n+1}\binom{n}{i-1}p^{i-1}q^{n-i+1}\log\frac{i(i-1)(n-i+1)}{(n+1)np^{i}q^{n-i+1}}\right) \\
& +q\left(\sum_{i=0}^{n}\binom{n}{i}p^{i}q^{n-i}\log\frac{(n+1-i)i!(n-i)!}{(n+1)n!p^{i}q^{n-i+1}}\right) \\
= \; & \left(p\sum_{i=1}^{n+1}\binom{n}{i-1}p^{i-1}q^{n-i+1} + q\sum_{i=0}^{n}\binom{n}{i}p^{i}q^{n-i}\right)\log\frac{1}{\binom{n+1}{i}p^{i}q^{n+1-i}} \\
= \; & -\left(q\binom{n}{0}q^{n} + \sum_{i=1}^{n}\left\{\binom{n}{i-1}+\binom{n}{i}\right\}p^{i}q^{n-i+1} + p\binom{n}{n}p^{n}\right) \\
& \times\log\binom{n+1}{i}p^{i}q^{n+1-i} \\
= \; & -\sum_{i=0}^{n+1}\binom{n+1}{i}p^{i}q^{n+1-i}\log\binom{n+1}{i}p^{i}q^{n+1-i} \\
= \; & H(P_{X_{n+1}})
\end{aligned}
$$

よって (8) の左辺になる. □

Corollary 1 (Jensen-Shannon divergence)

$$
\mathcal{JSD}(P,Q) \equiv \frac{1}{2}D\left(P||\frac{P+Q}{2}\right) + \frac{1}{2}D\left(Q||\frac{P+Q}{2}\right) \tag{9}
$$

$p = \frac{1}{2}$ のとき, $H(P_{X_n}) = H(P_{X_n+1})$ なることに注意すると (8) は (9) より

$$H(P_{X_{n+1}}) = H(P_{X_n}) + \mathcal{JSD}(P_{X_n}, P_{X_n+1})$$

となり, 二項分布のエントロピーは n に関して増加である. ここで, $\frac{Y_n}{n}$ が n に関して増加であることと次の式は同値である.

$$\log \frac{Y_{n+1}}{n+1} \geq \log \frac{Y_n}{n}$$

また $Y_n = e^{2H(X_n)}$ を用いると次のようになる.

$$\mathcal{JSD}(P_{X_n}, P_{X_n+1})) \geq \frac{1}{2} \log \frac{n+1}{n} \tag{10}$$

Lemma 2 任意の分布 P, Q と任意の $\alpha, \beta \geq 0(\alpha+\beta = 1)$ に対して, Jensen Shannon divergence は次のように拡張される.

$$\begin{aligned} C(P,Q) &\equiv \alpha D(P\|\alpha P + \beta Q) + \beta D(Q\|\alpha P + \beta Q) \\ &= \sum_{\nu=1}^{\infty} \frac{1}{2\nu(2\nu-1)} \Delta_\nu(P,Q) - \log 2 - \alpha \log \alpha - \beta \log \beta, \end{aligned}$$

ただし

$$\Delta_\nu(P,Q) = \sum_{i=1}^{n} \frac{|p_i - q_i|^{2\nu}}{(p_i + q_i)^{2\nu-1}}.$$

ここで $\alpha = \beta = \frac{1}{2}$ のときは

$$C(P,Q) = \sum_{\nu=1}^{\infty} \frac{1}{2\nu(2\nu-1)} \Delta_\nu(P,Q).$$

Proof. 次のようにおく.

$$\begin{aligned} m_i &= \alpha q_i + \beta p_i \\ \varepsilon_i &= |\alpha q_i - \beta p_i| \\ k_i &= \frac{m_i}{\varepsilon_i} = \frac{\alpha q_i + \beta p_i}{|\alpha q_i - \beta p_i|} \end{aligned}$$

このとき各 i について次の関係が成り立つ.

$$
\begin{aligned}
& \beta p \log \frac{p}{m} + \alpha q \log \frac{q}{m} && (11)\\
= \ & \beta p \log \frac{\beta p}{(\alpha q + \beta p)\beta} + \alpha q \log \frac{\alpha q}{(\alpha q + \beta p)\alpha} \\
= \ & \beta p \log \frac{\beta p}{\alpha q + \beta p} + \beta p \log \frac{1}{\beta} + \alpha q \log \frac{\alpha q}{\alpha q + \beta p} + \alpha q \log \frac{1}{\alpha} \\
= \ & \max\{\beta p, \alpha q\} \log \frac{\max\{\beta p, \alpha q\}}{\alpha q + \beta p} + \min\{\beta p, \alpha q\} \log \frac{\min\{\beta p, \alpha q\}}{\alpha q + \beta p} \\
& -\beta p \log \beta - \alpha q \log \alpha \\
= \ & \frac{m+\varepsilon}{2} \log \frac{1/2(m+\varepsilon)}{m} + \frac{m-\varepsilon}{2} \log \frac{1/2(m-\varepsilon)}{m} && (12)\\
& -\beta p \log \beta - \alpha q \log \alpha \\
= \ & \frac{\varepsilon}{2} \left(\frac{m}{\varepsilon} + 1 \right) \log \frac{1/2(m/\varepsilon + 1)}{m/\varepsilon} + \frac{\varepsilon}{2} \left(\frac{m}{\varepsilon} - 1 \right) \log \frac{1/2(m/\varepsilon - 1)}{m/\varepsilon} \\
& -\beta p \log \beta - \alpha q \log \alpha \\
= \ & \frac{\varepsilon}{2} \left\{ (k+1) \log \frac{1/2(k+1)}{k} + (k-1) \log \frac{1/2(k-1)}{k} \right\} \\
& -\beta p \log \beta - \alpha q \log \alpha \\
= \ & \frac{\varepsilon}{2} \left\{ k \log \left(1 - \frac{1}{k^2} \right) + \log \left(1 + \frac{1}{k} \right) - \log \left(1 - \frac{1}{k} \right) - 2k \log 2 \right\} \\
- \ & \beta p \log \beta - \alpha q \log \alpha && (13)
\end{aligned}
$$

したがって

$$
\begin{aligned}
C(P,Q) &= \sum_{i=0}^{n+1} \beta p_i \log \frac{p_i}{m_i} + \sum_{i=0}^{n+1} \alpha q_i \log \frac{q_i}{m_i} \\
&= \sum_{i=0}^{n+1} \frac{\varepsilon_i}{2} \left\{ k_i \log \left(1 - \frac{1}{k_i^2} \right) + \log \left(1 + \frac{1}{k_i} \right) - \log \left(1 - \frac{1}{k_i} \right) - (2 \log 2) k_i \right\}
\end{aligned}
$$

ここで (13) より

$$
\begin{aligned}
C(P,Q) &= -\beta \log \beta \sum_{i=0}^{n+1} p_i - \alpha \log \alpha \sum_{i=0}^{n+1} q_i \\
&= \sum_{i=0}^{n+1} \frac{\varepsilon_i}{2} \left(\sum_{\nu=1}^{\infty} \frac{1}{\nu(2\nu-1)} k_i^{-(2\nu-1)} - (2 \log 2) k_i \right) - \beta \log \beta - \alpha \log \alpha
\end{aligned}
$$

最後の式は $|k_i| < 1$ と $\log(1+x)$ のマクローリン展開から得られる. したがって

$$
C(P,Q) = \sum_{i=0}^{n+1} \varepsilon_i \sum_{\nu=1}^{\infty} \frac{1}{2\nu(2\nu-1)} k_i^{-(2\nu-1)} - \log 2 \sum_{i=0}^{n+1} \varepsilon_i k_i - \beta \log \beta - \alpha \log \alpha
$$

$$= \sum_{i=0}^{n+1} \varepsilon_i \sum_{\nu=1}^{\infty} \frac{1}{2\nu(2\nu-1)} k_i^{-(2\nu-1)} - \log 2 - \beta \log \beta - \alpha \log \alpha$$

なぜなら $\sum_{i=0}^{n+1} \varepsilon_i k_i = \sum_{i=0}^{n+1} m_i = 1$ であることを用いた．したがって

$$\begin{aligned} C(P,Q) &= \sum_{\nu=1}^{\infty} \frac{1}{2\nu(2\nu-1)} \sum_{i=0}^{n+1} \varepsilon_i k_i^{-(2\nu-1)} - \log 2 - \beta \log \beta - \alpha \log \alpha \\ &= \sum_{\nu=1}^{\infty} \frac{1}{2\nu(2\nu-1)} \Delta_\nu(P,Q) - \log 2 - \beta \log \beta - \alpha \log \alpha \end{aligned}$$

ただし

$$\Delta_\nu(P,Q) = \sum_{i=0}^{n+1} \varepsilon_i k_i^{-(2\nu-1)} = \sum_{i=0}^{n+1} \frac{\varepsilon_i^{2\nu}}{m_i^{2\nu-1}} = \sum_{i=0}^{n+1} \frac{|\alpha q_i - \beta p_i|^{2\nu}}{(\alpha q_i + \beta p_i)^{2\nu-1}}$$

とおいた． □

Lemma 2 を $P = P_{X_n}, Q = P_{X_n+1}, \alpha = q, \beta = p$ の特別な場合に用いると次の結果を得る．

Lemma 3 P_{X_n} と P_{X_n+1} について拡張された Jensen Shannon divergence は次のようになる．

$$\begin{aligned} C(P_{X_n}, P_{X_n+1}) &= \sum_{\nu=1}^{\infty} \frac{1}{\nu(2\nu-1)} \frac{2^{2\nu-1}}{(n+1)^{2\nu}} \sum_{i=0}^{n+1} \left(i - \frac{n+1}{2}\right)^{2\nu} (qp_i + qp_i) \\ &\quad - \log 2 - p \log p - q \log q \end{aligned}$$

ここで $p = q = \frac{1}{2}$ のときは $m_{2\nu}(B(n+1, \frac{1}{2}))$ を二項分布 $B(n+1, \frac{1}{2})$ の 2ν 次の中心積率 (central moment) とすると次のようになる．

$$C(P_{X_n}, P_{X_n+1}) = \sum_{\nu=1}^{\infty} \frac{1}{\nu(2\nu-1)} \frac{2^{2\nu-1}}{(n+1)^{2\nu}} m_{2\nu} \left(B\left(n+1, \frac{1}{2}\right)\right)$$

Proof. 次のことに注意する．

$$\begin{aligned} \Delta_\nu(P,Q) &= \sum_{i=0}^{n+1} \frac{|pq_i - qp_i|^{2\nu}}{(pq_i + qp_i)^{2\nu-1}} \\ &= \sum_{i=0}^{n+1} \left(\frac{|pq_i - qp_i|}{pq_i + qp_i}\right)^{2\nu} (pq_i + qp_i) \\ &= \sum_{i=0}^{n+1} \left(\frac{pq_i - qp_i}{pq_i + qp_i}\right)^{2\nu} (pq_i + qp_i) \end{aligned}$$

ここで

$$\begin{aligned}\frac{pq_i - qp_i}{pq_i + qp_i} &= \frac{p\binom{n}{i-1}p^{i-1}q^{n-i+1} - q\binom{n}{i}p^i q^{n-i}}{p\binom{n}{i-1}p^{i-1}q^{n-i+1} + q\binom{n}{i}p^i q^{n-i}} \\ &= \frac{i-(n-i+1)}{i+(n-i+1)} = \frac{2i-n-1}{n+1}\end{aligned}$$

したがって,

$$\begin{aligned}\Delta_\nu(P,Q) &= \sum_{i=0}^{n+1}\left(\frac{2i-n-1}{n+1}\right)^{2\nu}(pq_i+qp_i) \\ &= \left(\frac{2}{n+1}\right)^{2\nu}\sum_{i=0}^{n+1}\left(i-\frac{n+1}{2}\right)^{2\nu}(pq_i+qp_i)\end{aligned}$$

ここで $p=q=\frac{1}{2}$ のとき

$$\sum_{i=0}^{n+1}\left(i-\frac{n+1}{2}\right)^{2\nu}(pq_i+qp_i) = m_{2\nu}\left(B\left(n+1,\frac{1}{2}\right)\right)$$

□

4 離散型 Entropy Power 不等式

$n \curvearrowright \frac{Y_n}{n}$ が増加関数であることと (10) は同値である.

Theorem 1 $p=q=\frac{1}{2}$ のとき確率変数系列 X_n は以下の 離散型の Entropy Power 不等式をみたす.

$$^\forall m,n \geq 1,\quad e^{2H(X_m)} + e^{2H(X_n)} \leq e^{2H(X_m+X_n)} = e^{2H(X_{m+n})}$$

Proof. 次の不等式を示せば十分である.

$$\sum_{\nu=1}^{3}\frac{1}{\nu(2\nu-1)}\frac{2^{2\nu-1}}{(n+1)^{2\nu}}m_{2\nu}\left(B\left(n+1,\frac{1}{2}\right)\right) \geq \frac{1}{2}\log\left(1+\frac{1}{n}\right) \tag{14}$$

ただし $m_{2\nu}(B(n+1,\frac{1}{2}))$ の $\nu=1,2,3$ の中心積率はそれぞれ求めると次のようになる.

$$m_2 = \frac{n+1}{4},\ m_4 = \frac{(n+1)(3n+1)}{16},\ m_6 = \frac{(n+1)(15n^2+1)}{64}.$$

ここでこれらを (14) の左辺に代入して計算すると (14) が成り立つことがわかる. □

$p \neq \frac{1}{2}$ の場合, Mathematica の数値計算によると, 次の式のグラフが得られる.

$$\Delta_n = \frac{e^{2H(X_{n+1})}}{n+1} - \frac{e^{2H(X_n)}}{n} \tag{15}$$

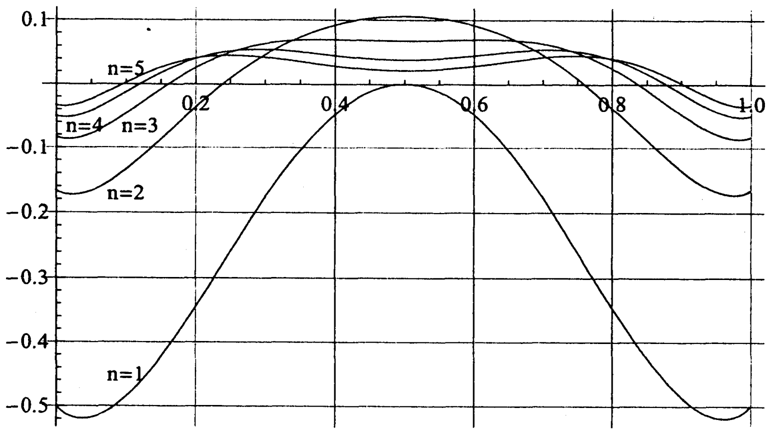

Figure 1: Δ_n のグラフ $(n = 1, 2, 3, 4, 5)$

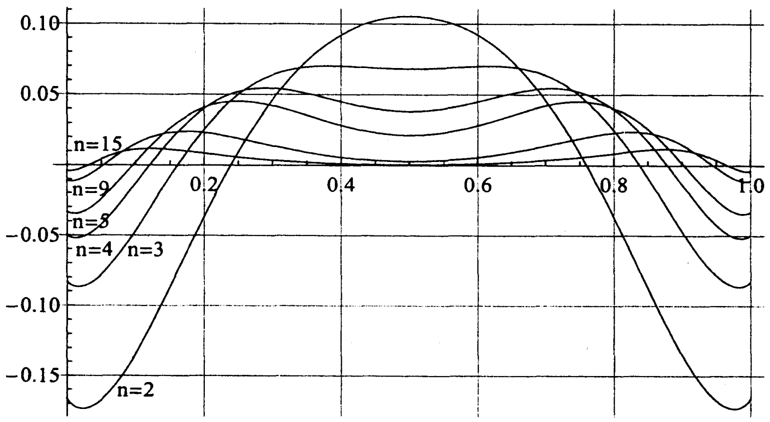

Figure 2: Δ_n のグラフ $(n = 2, 3, 4, 5, 9, 15)$

したがって次の Conjecture が考えられる.

Conjecture 1 $m, n \geq 2$ かつ $\frac{1}{4} \leq p \leq \frac{3}{4}$ のとき, 確率変数系列 X_n は以下の 離散型の Entropy Power 不等式をみたす.

$$^{\forall}m, n \geq 2, \quad e^{2H(X_m)} + e^{2H(X_n)} \leq e^{2H(X_m + X_n)} = e^{2H(X_{m+n})}?$$

この Conjecture が成り立つことを示すには, 次の不等式を示せばよいが, 今後の課題として残しておく.

$$\begin{aligned}
&\sum_{\nu=1}^{\infty} \frac{1}{2\nu(2\nu-1)} \left(\frac{2}{n+1}\right)^{2\nu} \sum_{i=0}^{n+1} \left(i - \frac{n+1}{2}\right)^{2\nu} (pq_i + qp_i) \\
&- \log 2 - p \log p - q \log q \\
\geq\ & \frac{1}{2} \log \left(1 + \frac{1}{n}\right)
\end{aligned}$$

References

[1] N.M.Blachman, The convolution inequality for entropy powers, IEEE Trans. Information Theory, vol IT-11, pp 267-271, 1965.

[2] M.H.M.Costa, A new entropy power inequality, IEEE Trans. Information Theory, vol IT-31, pp 751-760, 1085.

[3] A.Dembo, Simple proof of the concavity of the entropy power with respect to added Gaussian noise, IEEE Trans. Information Theory, vol IT-35, pp 887-888, 1989.

[4] P.Harremoës and C.Vignat, An entropy power inequality for the binomial family, J.Ineq.Pure and Appl.Math., vol 4, Issue 5, Article 93, 2003.

[5] S.Shamai and A.Wyner, A binary analog to the entropy-power inequality, IEEE Trans. Information Theory, vol IT-36, pp 1428-1430, 1990.

[6] C.E.Shannon, A mathematical theory of communication, Bell Syst. Tech. J., vol 27, pp 379-423 and pp 623-656, 1948.

[7] A.J.Stam, Some inequalities satisfied by the auantities of information of Fisher and Shannon, Information and Control, vol 2, pp 101-112, 1059.

[8] F.Topsoe, Some inequalities for information divergence and related measures of discrimination, IEEE Trans. Information Theory, vol IT-46, no 4, pp 1602-1609, 2000.